

# What are the social impacts of deepfakes?

## Abstract

Images and videos generated by computers, also called deepfakes, are usually created by automatically changing their facial attributes. They can have beneficial social applications and harmful effects. It is crucial to mitigate the negative consequences while allowing helpful applications.

**Rafael Luiz Testa**

## References

1. Albahar, M., & Almalki, J. (2019). Deepfakes: Threats and countermeasures systematic review. *Journal of Theoretical and Applied Information Technology*, 97(22), 3242-3250.
2. Chawla, R. (2019). Deepfakes: How a pervert shook the world. *International Journal of Advance Research and Development*, 4(6), 4-8.
3. Choi, Y., Choi, M., Kim, M., Ha, J. W., Kim, S., & Choo, J. (2018). Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 8789-8797).
4. Chesney, B., & Citron, D. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *Calif. L. Rev.*, 107, 1753.
5. Gosse, C., & Burkell, J. (2020). Politics and porn: how news media characterizes problems presented by deepfakes. *Critical Studies in Media Communication*, 37(5), 497-511.
6. Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys (CSUR)*, 54(1), 1-41.
7. Testa, R. L., Corrêa, C. G., Machado-Lima, A., & Nunes, F. L. (2019). Synthesis of facial expressions in photographs: characteristics, approaches, and challenges. *ACM Computing Surveys (CSUR)*, 51(6), 1-35.

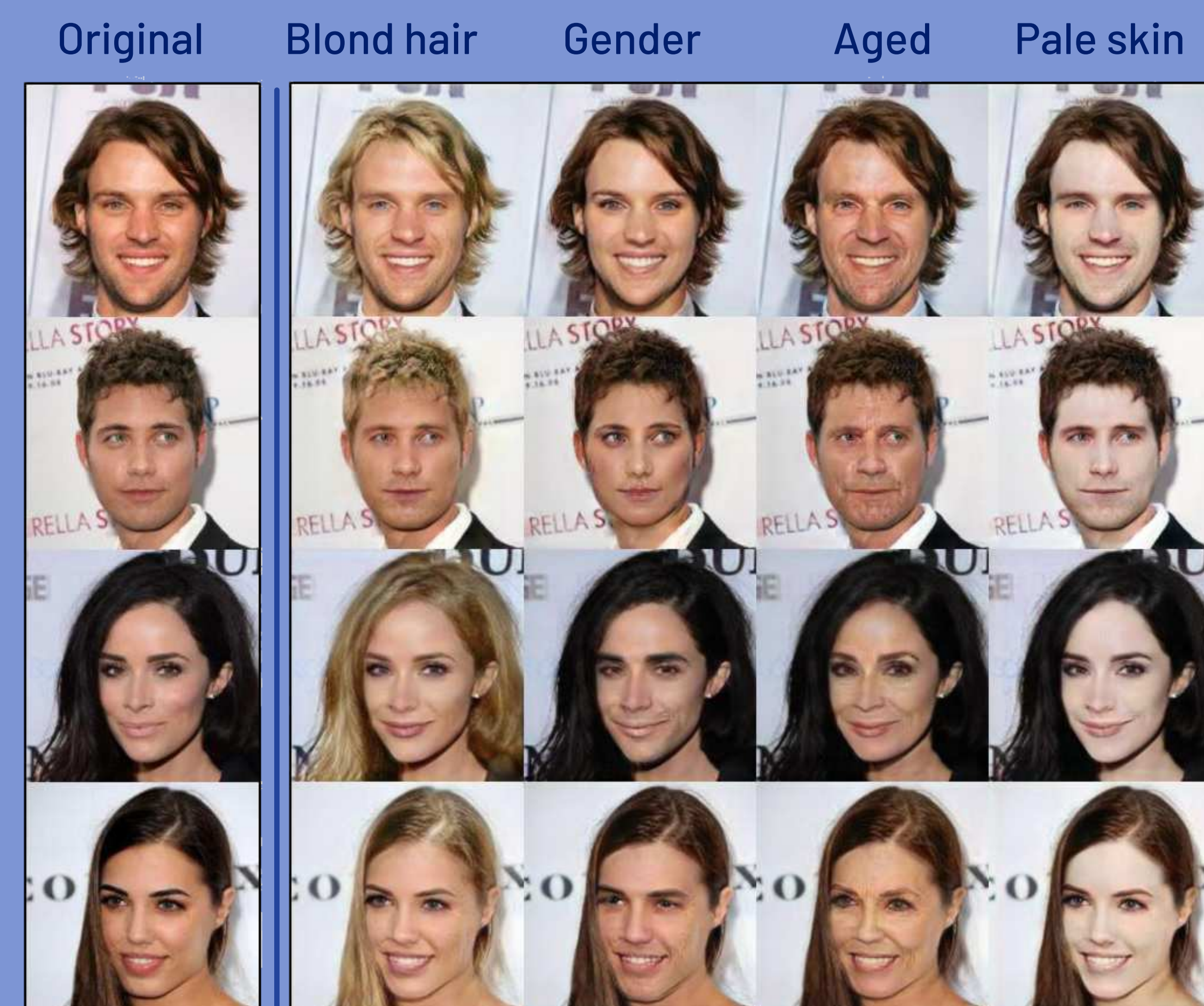
01

## Introduction

Deepfakes are a type of media generated by a computer that looks authentic to human perception.

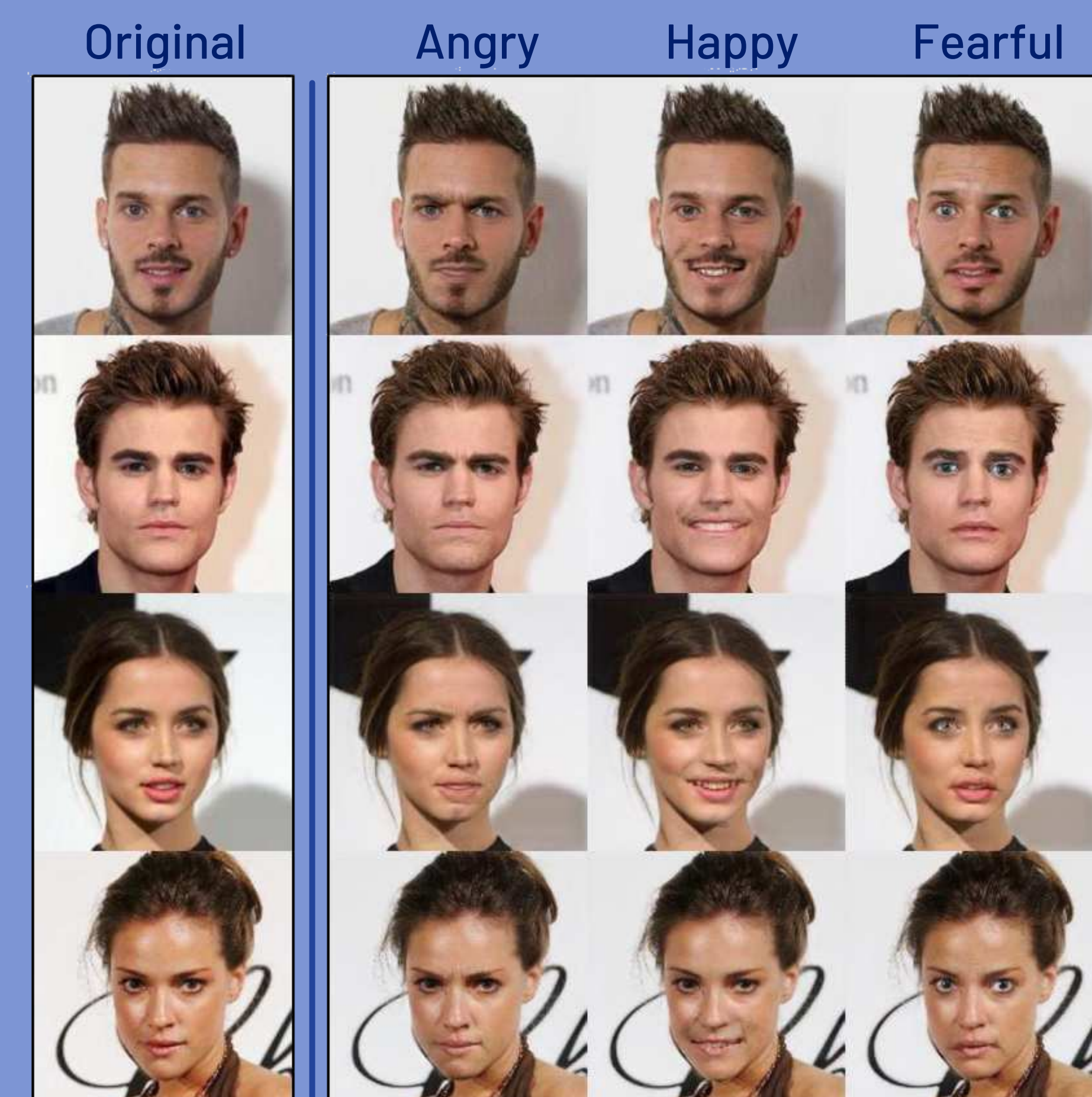
- Usually focus on facial images or videos (Mirsky & Lee, 2021).
- Changes attributes of the original image (Mirsky & Lee, 2021).

Figure 1. Facial attributes edited in images.



Adapted from Choi et al. (2018).

Figure 2. Facial expression edited in images.



Adapted from Choi et al. (2018).

Deepfakes can have beneficial and harmful impacts on society.

02

## Beneficial applications

- treatment and diagnosis of psychiatry disorders (Testa et. al, 2019);
- personalized media (Chawla, 2019);
- videos of historical figures (Chesney & Citron, 2018);
- customizable avatars (Chesney & Citron, 2018);
- humorous content (Chesney & Citron, 2018);

Figure 3. Illustrations related to beneficial applications.

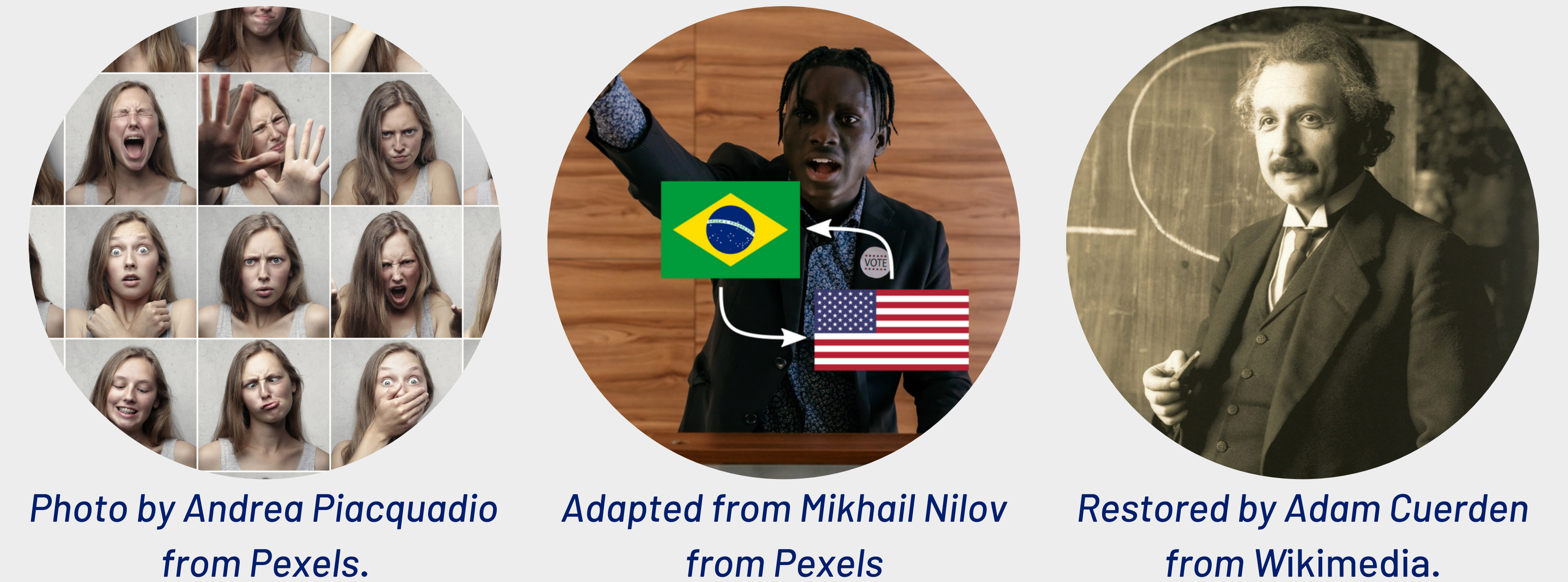


Photo by Andrea Piacquadio from Pexels.

Adapted from Mikhail Nilov from Pexels

Restored by Adam Cuerden from Wikimedia.

03

## Harmful implications

- non-consensual pornographic content (Chesney & Citron, 2018);
- identity theft (Chesney & Citron, 2018);
- exploitation and sabotage (Chesney & Citron, 2018);
- extortion (Chesney & Citron, 2018);
- question authentic content (Gosse & Burkell, 2020);

Figure 4. Illustrations related to harmful implications.



Photo by Andrea Piacquadio from Pexels.

Photo by Markus Winkler from Pexels.

Photo by Sora Shimazaki from Pexels.

04

## Conclusion

Detecting deepfakes may help to reduce the dissemination of fake content, which can mitigate its negative consequences. At the same time, deepfake videos have several applications that can result in benefits to society. The democratization of deepfake creation may lead to more awareness of malicious media. In the future, it will be difficult to distinguish between real and fake media. Thus, it is essential to research alternative solutions to validate real content.

Figure 5. Faces of actor used in deepfake videos.

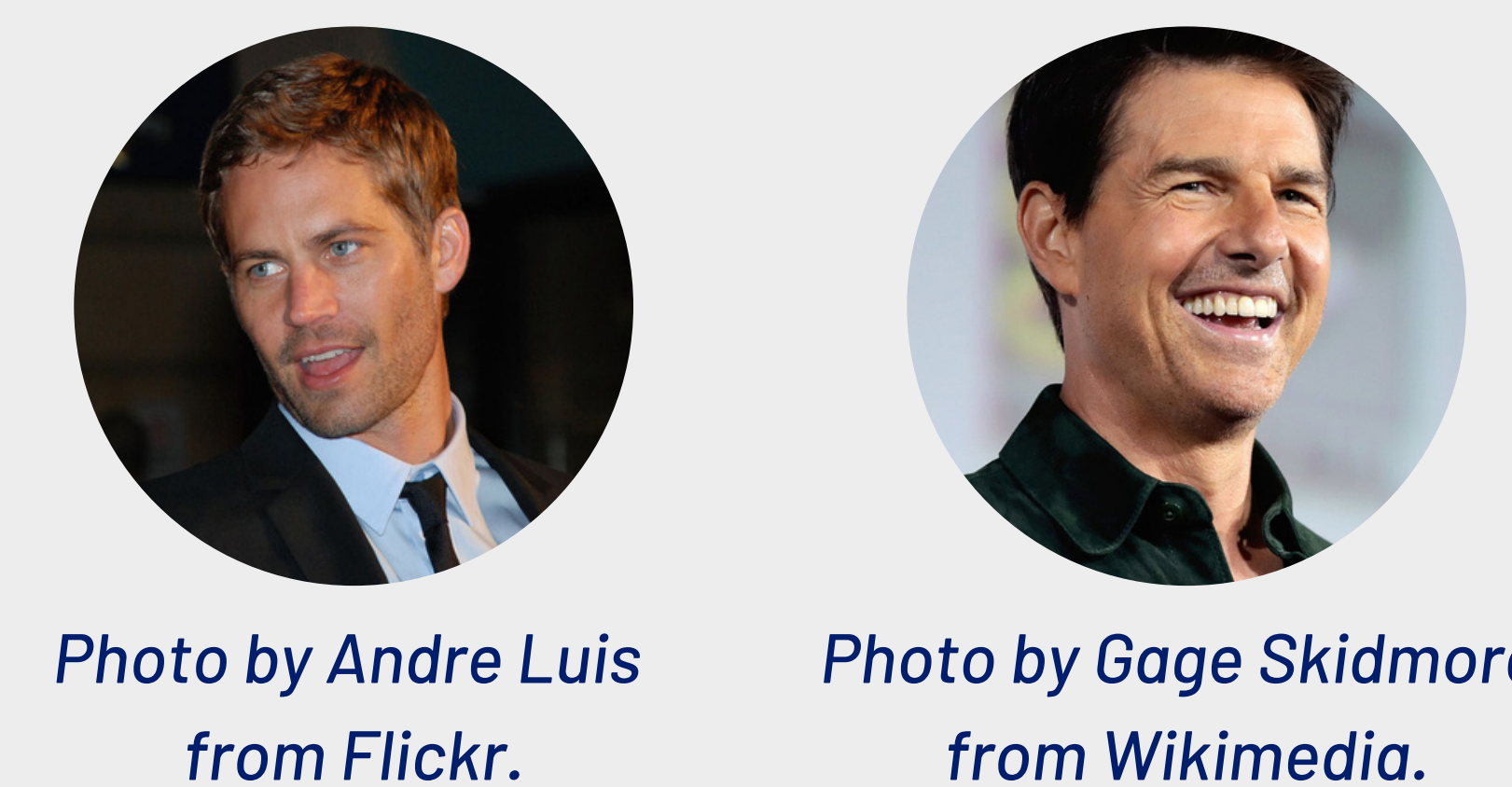


Photo by Andre Luis from Flickr.

Photo by Gage Skidmore from Wikimedia.